IMAGINATION ENGINES INC

Building synthetic brains capable of human level discovery and invention...

Company   **Concepts**   **Capabilities**   **Ventures**   **Articles**   **Patents**

IEI's Patented Creativity Machine® Paradigm

Summary - An artificial neural network that has been trained on some body of knowledge and then perturbed in a specially prescribed way tends to activate into concepts and/or strategies (e.g., new ideas) derived from that original body of knowledge. These transiently perturbed networks are called 'imagination engines' or 'imagitrons'. If another computational agent, such as a traditional rule-based algorithm or, even better, another trained neural network is allowed to filter for the very best of these emerging ideas, we arrive at an extremely valuable neural architecture, the patented Creativity Machine. Optional feedback connections between this latter computational agent and the imagination engine assure swift convergence toward useful ideas or strategies. ..This new AI paradigm is **vastly** more powerful than genetic algorithms (GA), efficiently generating new concepts on mere desktop computers rather than on the computational clusters required of GAs.

Details - Ordinary neural networks excel at learning from raw data, cumulatively learning to associate one pattern with another, as when raw sensory inputs from our five senses activate mental images or feelings. Note, however, that such a direct link between the external world, and our own internal mental life is only a small part of brain activity. To produce a more faithful emulation of human cognition, some mechanism must be established that provides an internal genesis of thoughts and ideas that draws upon cumulative experience, rather than what our senses are telling us at the moment.

Creativity Machines represent a new kind of neural network paradigm that is capable of generating rather than just associating patterns. They are based upon what we believe to be a significant scientific discovery: that a neural network exposed to any conceptual space and then internally irritated, in a specially prescribed way, tends to generate coherent ideas derived from its absorbed wisdom. Another neural network, trained to filter out the very best of these notions, patrols the former net's outputs, accumulating useful concepts (i.e., new drug or automobile designs) or using these output patterns in real time to devise strategies (i.e., robotics and control systems).

Perhaps the most appropriate benchmark for Creativity Machines is the highly popularized genetic algorithm (GA) wherein mathematically simulated 'genes' randomly combine and mutate to produce new potential offspring that in turn represent new concepts, designs, or strategies. The problem with such systems is that they work on geologic time scales relative to the inherent processing speed of digital computers. Furthermore, GAs fail to deliver solutions as the dimensionality of the problem increases and the algorithm's generative process leads to combinatorial explosion and nonsensical offspring.

Problem dimensionality is not an issue for Creativity Machines. In a nutshell, this advantage stems from the fact that this neural architecture emulates the thalamocortical loop of the brain (i.e., the seat of intelligence and consciousness) rather than blind, and excruciatingly slow processes of mutation and natural selection. Furthermore, the actual time to build a Creativity Machine is negligible to genetic programs, since the CM is a self-organizing system. GA's, in contrast are not, and require that human experts hardwire myriad constraint relationships into their human-originated computer codes. In other words, genetic algorithms must be written by human programmers, whereas Creativity Machines build themselves!

The benefit to you, as an IEI customer, is that a specially tailored Creativity Machine can generate results at typically a tenth of the cost of any other AI paradigm perceived as competitive. Furthermore, it can deal with complex problems that the other approaches can't even touch.

References

Yam, Philip (1993). "Daisy, Daisy: do computers have near-death experience?", *Scientific American*, May 1993.

Yam, Philip(1995). "As They Lay Dying ... Near the end, artificial neural networks become creative", *Scientific American*, May, 1995.

Thaler, S. L. (1996). "Neural Networks That Create and Discover," *PC AI*, May/June 1996.
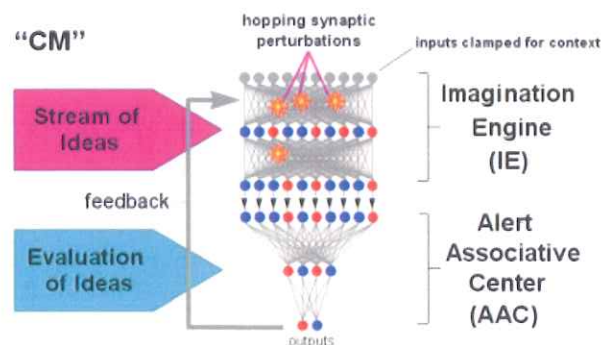
Holmes, R. (1996). "The Creativity Machine," *New Scientist*, 20 January 1996.

Thaler, S. L. (1997). US Patent 5,659,666, "Device for the autonomous generation of useful information", Issued August 19, 1997.

Thaler, S. L. (1997). US Patent 6,018,727, "Device for the autonomous generation of useful information", Issued January 25, 2000.

Thaler, S. L. (2000). US Patent 6,115,701, "Device for the autonomous generation of useful information", Issued September 5, 2000.
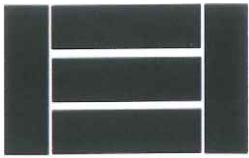
Thaler, S. L. (2002). US Patent 6,356,884, "System for the Autonomous Generation of Useful Information", Issued March 12, 2002.

Thaler, S. L. (2008). US Patent 7,454,388, "Device for the autonomous bootstrapping of useful information", Issued November 18, 2008.

Cohen, A. (2009), "Stephen Thaler's Imagination Engines," World Future Society, July 9, 2009.

IMAGINATION ENGINES INC

Building synthetic brains capable of human level discovery and invention...

Company   Concepts   Capabilities   Ventures   Articles   Patents

A Quantitative Model of Seminal Cognition: The Creativity Machine Paradigm (US Patent 5,659,666)

© 1997, Stephen L. Thaler, sthaler@imagination-engines.com, Imagination Engines, Inc., St. Louis, MO, USA, http://www.imagination-engines.com/, Presented to the Mind II Conference, Dublin, Ireland, 1997

Abstract - A synaptically perturbed neural network forms an efficient search engine within and around any conceptual space upon which it has been trained. By monitoring the temporal distribution of concepts emerging from such a system, we discover a quantitative agreement with the measured rhythm of human cognition, creative or otherwise. Closer examination of this transparent connectionist search engine suggests that much of human creativity may be attributed to the failure of cortical networks to activate into known memories as these networks perform vector completion upon their own internal disturbances. In lieu of intact memory activation, the networks produce a stream of degraded memories, now constituting what we commonly refer to as "ideas," that are filtered for utility and interest by attendant cortical networks.

Introduction

Creativity is perhaps the most celebrated of human capacities, embraced by the human potential movement and revered in the same light as other "folk" attributes such as *spirit, soul,* and *free will*. In the objective analysis of creativity, however, we must recognize that much of the grandeur and mystique of this cognitive phenomenon may be no more than a societal judgment that falls far short of established scientific standards. No longer squinting at the reality, we must account for why human progress is so desultory and why human intellectual activity does not take the most direct deductive path toward a final and ultimate product. Adhering to a reductionist model, we must account for ostensibly breathtaking paradigm shifts and innovations based upon a system of cortical neurons exchanging nothing more than matter and energy with the environment.

In taking this purely physical tack we must realize that just like swinging doors and molecules, the brain is a dynamical system endowed with various degrees of freedom. For a door the single degree of freedom describing its state is the rotation angle about its hinge axis. For a polyatomic molecule the degrees of freedom include the many allowed vibrational and rotational modes into which all conceivable motions of constituent atoms may be resolved. For the brain the allowed states are its neural activation patterns, each one of which represents some memory, sensation, or idea. We note that just as in the simpler physical systems, the total number of possible degrees of brain freedom is finite and attributable largely to the existence of electrochemical constraints (i.e., long-term potentiation) that bar the arbitrary activation of any given neuron or group of neurons within this sizable collective system.

Recognizing that any of these dynamical systems can only evolve in a manner dictated by internal constraints, we anticipate that the introduction of any random perturbation to these systems will drive them only through their allowed degrees of freedom. Therefore, the door will only move through its hinge axis when subjected to random jarring. A water molecule will respond to impact with other molecules by executing only its allowed translational, vibrational, and rotational modes. Likewise, when bathed in internal chaos, the cortex can only move through its allowed manifold of electrochemical states, each encompassing some idea or notion, whether mundane or profound. The result is what we commonly term *stream of consciousness*, a succession of thoughts, apparently from out of the blue.

Obviously, as long as the systems of electrochemical constraints in the brain remain intact, there can be no more seminal thought than the turnover of preexisting memories. One avenue toward creating new and unique activation patterns, and hence original ideas, is to destroy electrochemical constraint relations within cortical networks. Theoretically, exercising this single option opens the door not only to an intriguing model of creative cognition but also to a powerful computational paradigm.

The Creativity Machine Paradigm

Rumelhart and colleagues (1986) emphasized the utility of parallel distributed processing systems as constraint satisfaction networks in their pioneering work. Using "hand-wired" competitive networks exposed to various room schemata, they were able to demonstrate a primitive brand of creativity in which novel, yet plausible furniture combinations were predicted. Using the well-known principle of *vector completion*, the net could accept incomplete inputs representing a partially described room and through subsequent annealing could arrive at a fuller description of that room. Therefore, when only supplied with the inputs of a stove and a coffee cup, the net could finally arrive at a network state in which additional processing units corresponding to a refrigerator, sink, and oven could be likewise activated. In other words, the net was prescribing a plausible room setting that it may not have "seen" within its training experience. In this sense, such a network was *inventing* new room types.
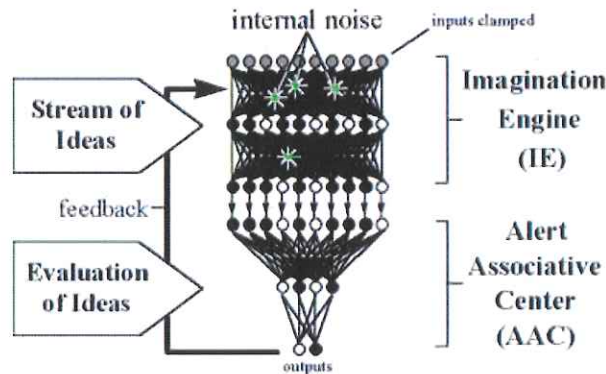
Figure 1. A Simple Creativity Machine. Here, the starbursts represent 'hopping' perturbations among the connection weights of the imagination engine.

Recently I have demonstrated (Thaler, 1995, 1996 a, b, c) that a trained artificial neural network supplied no inputs whatsoever, and driven by stochastic perturbations to its internal architecture may generate self-consistent schemata related to the conceptual space embodied within its training exemplars. In short, the network is perceiving something when in fact there are no presented environmental inputs. Accordingly I have coined the term "virtual input effect" to describe the phenomenon. Contacting with Rumelhart's work, if we were to train a simple auto-associative feedforward net on numerous examples of room schemata (hence bypassing the tedious Bayesian statistics used to construct his net), setting the inputs of the network to values of zero and then randomly perturbing the connection weights from their trained values, we would observe a progression of network activations corresponding to plausible furniture schemes. The difference in operating procedure from Rumelhart's work is significant, representing the distinction between *perception* with its processing of environmental features, and *internal imagery* with its inherent independence from such external entities. In the Rumelhart's original work, an associative net is interpreting some partial environmental vector as something it has never seen. In the case of virtual input effect, the net is in a state tantamount to sensory deprivation, in effect hallucinating within a silent and darkened room.

When supplied no external inputs, the production of meaningful activations by the network relies upon a different brand of vector completion than is normally discussed. Rather than fill in incomplete or corrupted input patterns, the net attempts to complete internal, noise-induced activation patterns within the net's encryption layers. Therefore, any local or temporary damage to the network's mapping is interpreted by downstream layers as some "familiar" activation pattern normally encountered upon application of a training exemplar to the network's inputs (Thaler, 1995). Because of the many combinatorial possibilities in perturbing connection weights within a network, we arrive at a means for generating proportionately more novel schema than is possible with input perturbations alone. Furthermore, because the connection traces within a trained neural network generally correspond to the rules binding the underlying conceptual space together, such stochastic perturbation schemes serve to soften these rules, in turn allowing a gradual departure from the known space of possibilities. The result is a strictly neurological search engine whose internal noise level may be parametrically increased to achieve progressively more novel concepts. I call such a chaotic network an *imagination engine* or IE.

By attaching to the IE a critic network (termed an *alert associative center* or AAC) that has been trained by example to recognize any emerging concept that possesses utility or value, a *Creativity Machine* is formed. Because the only inputs to this closed loop system take the form of unintelligible stochastic perturbations (i.e., heat), the system is deemed autonomous. Therefore, it monitors its own chaotically generated stream of consciousness, if you will, periodically extracting and isolating any concepts offering usefulness. The critic net may in turn modulate the intensity of perturbation within the first net, willfully dropping the computational temperature within the IE when that network appears to be on the right track (i.e., an attentional mechanism).

Table 1. Some Recent Creativity Machine Successes (1996)

Application Area	Outcome	Reference
musical composition	copyrighting of 11,000 novel musical 'hooks'	U.S. Copyright PAu-1-920-845 "Musical Themes From Creativity Machine"
materials discovery	autonomous generation of a materials database, including potentially new ultrahard materials and high-temperature	Autonomous Materials Discovery via Spreadsheet-Implemented Neural Network Cascades, JOM-e, 49(4) (1997) http://www.tms.org /pubs/journals /JOM/9704/Thaler

	superconductors	
beverage invention	a dynamic database of over 15,000 mixed drinks	http://www.imagination-engines.com/NeuralBar/Nbar.htm
personal hygiene product design	20% improvement in performance over existing designs	anonymous U.S. Corporation
control system	successful construction and testing of thin film coating reactor control system that invents recovery paths.	U.S. Air Force SBIR contract AF96-152, Automated Data Acquisition For In Situ Material Process Modeling

The practicality and successes (see Table 1 for a few examples) of the Creativity Machine paradigm stem from the fact that all networks involved are trained by example. Therefore, as long as historical data exist within any conceptual space, backpropagation or any other neural network learning paradigm may be used to rapidly train the required Creativity Machine networks. This ease of construction has allowed the building of a wide variety of Creativity Machines focused on diverse knowledge spaces, ranging from music composition, to ultrahard materials discovery, to the invention of personal hygiene products.

Common to the operation of most Creativity Machines built to date is a perturbation scheme in which small disturbances stochastically "hop" among the connection weights of the network. To parametrize the internal chaos within the IE, the governing algorithm parcels out n perturbations, usually of fixed or average magnitude s, then randomly and cyclically distributes them among the N total connection weights of the IE. In Figure 2, for instance, when network inputs are clamped, the governing algorithm places four perturbations (represented by starbursts) of fixed magnitude at time t_0, resulting in a distinct activation pattern at the network's outputs that represents some idea or concept. On every half cycle, $t_0 + dt/2$, the perturbations are removed, restoring the net to its trained-in state. Finally, in initiating a new cycle at time $t_0 + dt$, the algorithm randomly places the n perturbations of magnitude s on newly chosen connection weights. When viewed as a rapid graphical succession, the hopping motion resembles a boiling liquid, hence suggesting the term "cavitation" to describe this specific agenda of stochastic network perturbation.
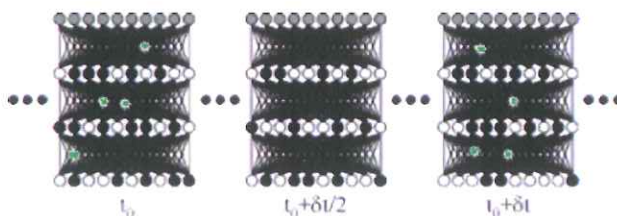

Figure 2. "Cavitation" of the Imagination Engine.

Therefore, during operation the Creativity Machine may be run under a whole range of operating conditions governed by the parameters n, s, and dt that collectively specify the level of cavitation applied to the IE. Obviously, applying no perturbation at all ($n = 0$ or $s = 0$) to the IE will result in no activation turnover and hence no idea generation. Alternatively, applying large perturbations n and s will produce such significant degradation to the network mapping that all constraints are destroyed within the captured knowledge domain. The result of severe perturbation is therefore to produce totally unconstrained activation patterns containing little, if any, information content. The former regime consists of vanishingly small perturbations and is regarded as "Neo-Lamarckian" in nature (Rowe and Partridge, 1993), representing a highly constrained and hence inefficient discovery mechanism. The latter unconstrained search regime, at high values of n and s, is considered "Neo-Darwinian" and is likewise inefficient due to the extensive sifting required by the critic network to find meaningful information among the multitudes of unconstrained concepts produced.

Obviously the ideal regime for Creativity Machine operation lies somewhere between the Neo-Darwinian and Neo-Lamarckian search regimes. To achieve the necessary level of internal perturbation, the parameters n and s are adjusted so that the quantity ns/N (where N is the total number of connection weights in the IE) is approximately 0.05–0.06, representing the mean perturbation per connection weight in the IE. Dividing through by dt, the perturbation time constant depicted in Figure 2, we obtain a parameter called the "cavitation rate,"

$$r = ns/Ndt, \qquad (1)$$

representing the mean rate of perturbation for any connection weight in the IE and the primary controlling parameter behind the

imagination engine.

Analytically, the choice of mean perturbation $ns/N = 0.06$ generally defines a cusp in network behavior that separates a regime of perturbation level corresponding to intact memory recall from that of increasingly corrupted memory generation (i.e., confabulation). This transition in the fidelity of network activations is a generally observed pattern among all IEs used to date and is exemplified in Figure 3, where we see this behavioral transition in a plot of the probability of intact memory production versus *cavitation rate* within a small internally perturbed network with constant inputs. The net has been trained to contain the memory of 16 binary vectors. This distinct separation between intact memory and confabulation persists even within more abstract conceptual spaces that may include subjective areas such as musical composition or more objective problems as in the discovery of new high-temperature superconductors (as discussed in Figure 4).
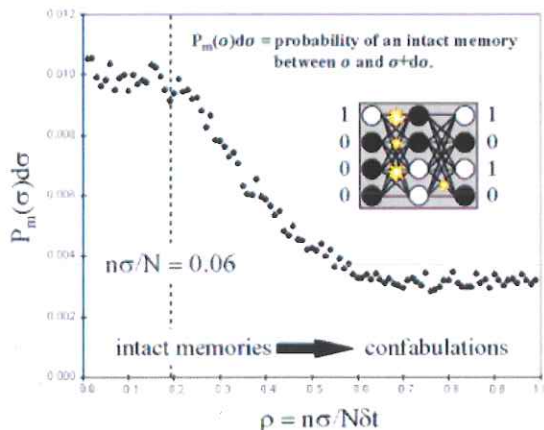

Figure 3. The Probability of the Noise-Induced Activation of an Intact Network Memory as a Function of the Cavitation Rate, r. Note the cusp near ns/N=0.06 dividing intact from corrupted memory recall. The plot is the result of 1,000 cavitation cycles applied to the simple auto-associative net shown in the inset, trained on 16 binary vectors, subjected to n=4 perturbations of variable magnitude s and a time constant dt of 0.3 sec. Inputs of the net were clamped at the binary memory (1,0,0,0).
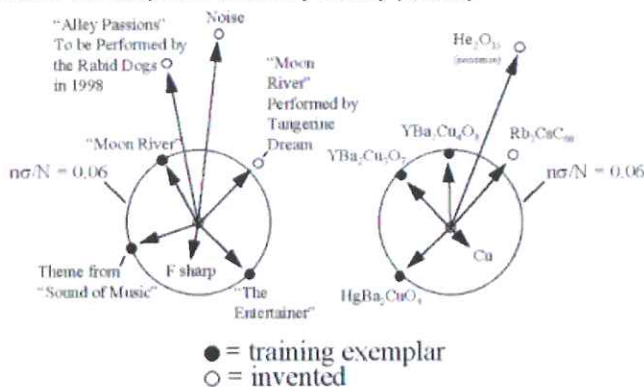

Figure 4. The invention of a plausible concept by the imagination engine takes place within a membrane surrounding the ns/N = 0.06 surface, corresponding to the cusp region in Figure 3. Excursions in mean connection weight perturbation significantly beyond this regime produce noise.

We find in general that the most fertile cavitation regime corresponds to mean connection weight perturbations near 0.06. At lower perturbation levels the IE revisits largely training exemplars and their generalizations. At progressively higher levels of connection weight perturbation, the IE produces less constrained and hence more nonsensical possibilities (i.e., noise).

Realizing that the connection weights of a neural network implicitly contain the rules and schema that bind together any given conceptual space, the perturbation scheme embodied within cavitation effectively experiments with these rules by softening them either individually or in parallel while the AAC judges the utility of the resulting concepts. A mean connection weight perturbation of approximately 0.06 appears to be a universal amount by which to soften these internal rules without producing nonsensical or known concepts. Symbolically representing the constraint relations within any given neural network as the unit sphere, coherent concepts that embody most of the useful ideas emerging from an IE fall within a thin membrane surrounding the *ns/N* = 0.06 surface, no matter what the conceptual space involved. Excursions too far beyond this surface, where *ns/N* >> 0.06, generally produce nonsense, as intimated in Figure 4.

The Choice of Objective Observables in the Scientific Modeling of Creative Cognition

Qualitatively, the Creativity Machine constitutes a compelling model of how both novel and mundane concepts may both nucleate within any parallel distributed system. Accordingly, it represents a strongly competitive functional metaphor for how the similarly connectionist brain creates. To search for a closer equivalence between the two systems, it is necessary to establish quantifiable

observables and then compare them. Having reduced the description to the most canonical level, we sidestep any top-down subjective description that is generally embellished by human cortical networks.

Within the quantitative dynamical system analogy, physics routinely recruits conjugate observables, such as position and momentum or energy and time, to describe the evolution of almost everything. Assuming no special physical status for the brain, its behavior may be adequately described by similar quantities describable by phase spaces whose axes correspond to these conjugate quantities. In the most rigorous of portrayals, cortical activation patterns would be described by a huge multidimensional space whose axes would correspond to the on–off status of each of the roughly 100 billion cortical neurons, along with a similarly immense space of rates of change of each of these neuron activations. In essence, this description represents a dichotomy, distinguishing what is being thought (i.e., the exact activation pattern in the former subspace) and how those activations are evolving in time (the latter subspace). Presently surrendering all hope of understanding exactly what is being thought, we may readily monitor and model the temporal pattern of cognitive turnover.

Accordingly, the first clue of a temporal link between the Creativity Machine and human cognition comes when an audible tone is attached to each Creativity Machine discovery. Listening to the resulting stream of alarms one detects a clustered distribution, with discoveries generally clumped together. Run at high noise levels, the stream attains the rhythm or prosody (Kosslyn and Koenig, 1992) of human speech, sounding much like a garbled conversation. To quantitatively examine the suggested correspondence to the human cognitive rhythm, we first calculate temporal distributions experimentally for both the Creativity Machine and cognitive streams for human test subjects. The measured temporal behaviors of both neurobiological and computational neural systems are shown to be identical, with both turnover rates derivable from the theory of fractal Brownian motion (fBm).

Measurement of Concept Generation Prosody Through Mandelbrot Measures

Intuitively we are well aware of the fact that the temporal distribution of thought shows similar clustering behavior over different temporal regimes. For instance, the musical output of a great composer may show a clustering over time, consisting of lull periods of inactivity peppered with spasms of creative turnover over months or years (Jamison, 1994, in the context of manic-depressive illness). Within the course of a single day, that composer's musical output may display similar surges and lulls. Likewise, in speaking we tend to produce a grouping of words as some main theme or idea appears to us followed by a noticeable lag as a new train of thought emerges. Similar clustering then appears at the level of sentences and individual words within those sentences. Therefore, to arrive at some convention for measuring temporal distribution, we not only require some means to measure the probability that any thought will accompany any other thought within a given time frame, but we also require some measure of any time-scale invariance involved. The natural way of approaching this problem is in the context of fractal theory where we are accustomed to examining spatial invariance (i.e., the coastline of Great Britain at various levels of magnification, where a satellite view is statistically indistinguishable from a view from several feet).
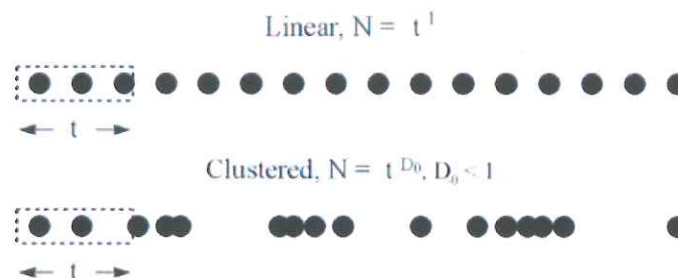

Figure 5. Fractal Dimension Calculation for Two Distinctive Temporal Distributions.

Consider for instance the generic temporal stream of events pictured at the top of Figure 5 where we see a distribution of equally separated events occurring at regular intervals. Randomly moving statistical sampling boxes of different durations t over the distribution we will find that the average number of captured events scales as t^1. Because of the unitary exponent, we say that this distribution possesses both a Euclidean and a fractal dimension of 1. In contrast, the lower event stream of Figure 5 depicts a nonlinear distribution that yields a fractal dimension of less than 1 through the same statistical sampling process. Generally in fractal studies the fractal dimension D is determined by what has generally become known as Mandelbrot measures (Mandelbrot and van Ness, 1982).

In the Mandelbrot analysis, P(m, t) is defined as the probability of statistically measuring m points within a sampling time of duration t. A computer code may calculate this quantity by "dropping" sampling boxes of progressively larger time frames t onto the resulting distribution and then counting the number of bracketed events. For each sampling box of time t, the algorithm may perform multiple random samplings of the distribution. P(m, t) is then normalized such that

$$\sum_{m=1}^{N} P(m,t) = 1, \qquad (1)$$

for all t, where N is the total number of points within the sampled system. The distribution P(m,t) is then used to define the mass moments

$$M^q(t) = \sum_{m=1}^{N} m^q P(m,t), \qquad (2)$$

where q assumes positive integer values. The fractal dimension, D, is then estimated from the logarithmic derivatives,

$$D_q = \frac{1}{q} \frac{\partial \log [M^q(t)]}{\partial \log [t]}, \qquad (3)$$

or from linear plots as shown in Figure 3. Generally, if the same fractal dimension D_q applies to a range of q values, we say that a fractal interpretation applies over a range of event cluster sizes. If the calculated fractal dimension is identical for values of q=1 to 3 (as in Figure 6), then individual events are statistically distributed in the same way as clusters containing 2 or 3 events.
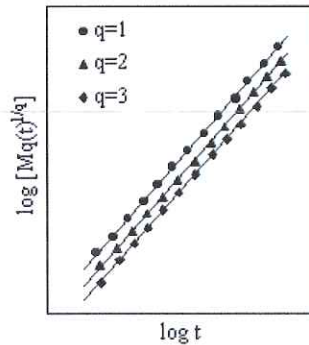

Figure 6. Graphical determination of the fractal dimension from the slope of the logarithm of mass moment versus the logarithm of time, t.

Mandelbrot Analysis Applied to Cognitive Experiments

Twelve volunteers contacted by telephone were asked to name 20 items as quickly as possible for each of the series of topics listed in Table 2, while digitally recording their responses. Test subjects thereby tacitly assumed that the objective of the experiment was to note speed, rather than the sought distribution of their thought stream. The desired effect was then to minimize the latent period between idea formation and concept articulation to approximate as closely as possible the arrival times of consecutive thoughts. The resulting digital strip-chart recording (as exemplified in Figure 7) was then used to quantify the cognitive event stream by noting the start of each word or phrase on a millisecond time scale. Stuttering, which was rare within this study, was considered the leading edge to any voiced concept. Subsequent analysis of this event stream, by the methods of Mandelbrot analysis, yielded both a total observation time Dt and a fractal dimension D. The combination of the fractal characteristics along with the total time scale required to complete the cognitive task constituted a complete statistical, temporal description of the cognitive event stream. Because the calculated fractal dimension is intrinsically time-scale independent and measured total time Dt is explicitly time-scale dependent, the two parameters form a complementary set of temporal observables.
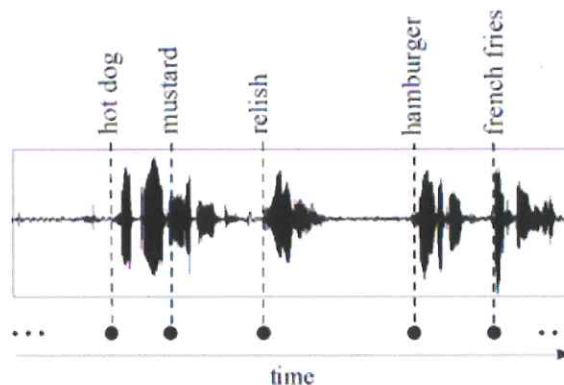

Figure 7. A Representative Strip Chart Recording for the "Foods" Cognitive Task.

In retrospect, it is noteworthy to mention that within the context of these cognitive experiments, the resulting fractal calculations appeared independent of any foreknowledge by the test subject of the intent of the experiment. This observation may be testament to inability of human cognitive faculties to store a large number of thoughts while simultaneously counterfeiting a bogus pattern of articulation. Only in well-rehearsed cases or in reading from written lists could test subjects attain arbitrary speech rhythms.

Table 2. Topic Areas for the Cognitive Tasks

Rote Cognitive Tasks	Creative Cognitive Tasks
Name 20 numbers.	Invent 20 nonsense words beginning with the letter "r."
Name 20 foods.	Name 20 ways to enter a house.
Name 20 Mexican foods.	
Name 20 states.	
Name 20 words beginning with the letter "r."	

Mandelbrot Analysis of Creativity Machine Activation Turnover

An auto-associative network was trained by the standard methods of backpropagation ([Rumelhart, Hinton, and Williams, 1986](#)) by exposing it to identical binary input–output vectors. The network contained four layers of 10 processing units each, all fully connected between layers, for a total of 300 connection weights. Input and output training exemplars for the network consisted of generic binary memories, ranging from (0, 0, 0, 0, 0, 0, 0, 0, 0, 0) to (1, 1, 1, 1, 1, 1, 1, 1, 1, 1). Once trained, this net was embedded within a C code that cyclically supplied perturbations of fixed magnitude s to n randomly chosen connection weights, as described above and depicted in Figure 2.

To normalize the output of this network to that of the cognitive study, the perturbation time constant dt was adjusted to a value of 300 msec to correspond to the fastest enunciation rate of ideas (i.e., the recall of 20 numbers, invariably in ordinal fashion). Therefore, the maximum rate at which the network could produce its own activation turnover was adjusted to correspond to the fastest possible cognitive rate measured experimentally. So normalized, the net was run at the optimal mean connection weight perturbation of 0.057, yielding a cavitation rate of $r = 0.057/0.3 \text{ sec} = 0.19 \text{ sec}^{-1}$. A variety of network perturbation schemes (i.e., different combinations of n and s) were used so that the cavitation rate of 0.19 sec^{-1} was maintained constant.

Following each distinct redistribution of perturbations among the N weights, fixed inputs of 1/2 were fed through the network with the controlling algorithm noting whether a transition (i.e., 50% change in any output vector component) had occurred. These transitions then constituted the activation turnover of the net. Simulation halted after 20 such transitions, at which time the algorithm applied fractal dimensional analysis to the recorded network output transitions using the same algorithm to determine Mandelbrot measures as was used in the cognitive study.

Comparison of Cognitive and Creativity Machine Prosodies

Having amassed roughly 100 cognitive experiments and a similar number of experiments on IE activation turnover, a number of curve fitting experiments were carried out to investigate whether any simple pattern existed between the calculated fractal dimension and measured time scales for either data base. For both sets of experiments it was found that D_0, the calculated fractal dimension for either event stream, was inversely proportional to the logarithm of the total time scale required to complete the task. In Figure 8 we see the striking similarity between the two characterizations. We find that the full temporal characterizations are equivalent to within the experimental error of the study.

Empirically, we find that both cognition and the chaotic ANN both obey the same law of event turnover given by

$$D_0 = 1.62 / \ln(Dt), \qquad (4)$$

where Dt is expressed in seconds and the proportionality constant represents the mean between the cognitive and ANN result of Figure 8 and all previous studies (Thaler, 1996b). Recasting Equation 2 into exponential form, we obtain the relation

$$Dt^{-D_0} = 0.19, \qquad (5)$$

indicating that for both the sampled cognition and the Creativity Machine, there exists a trade-off between the time scale required for a set number of distinct transitions and the inherent clustering of events within the resulting transition sequence.
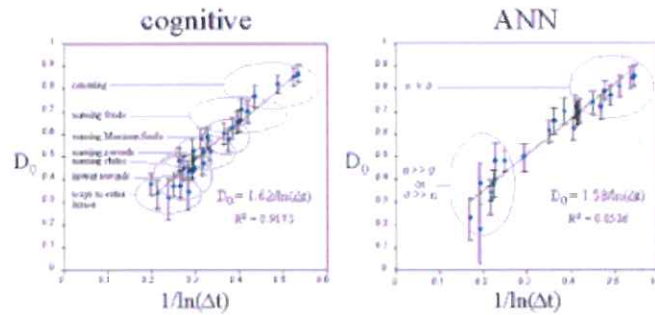
Figure 8. Full Temporal Characterization of Both Cognitive and Creativity Machine (ANN) Event Streams. Cognitive data points represent as many as 5 repetitions of the same cognitive task. The ANN data points represent 5 repetitions of the same computer experiment. Dt is in units of seconds.

In the case of human speech, this result is intuitively familiar: a speaker who is familiar with his or her presentation material tends to speak in a relatively linear fashion. In contrast, with more ad lib delivery the speaker's articulated words tend to be more clustered. In this sense, the relationship embodied in Equation 5 is a quantitative expression of this all too familiar phenomenon of hesitancy.

In the plot for the cognitive study, some of the specific tasks were overlaid to display their relative positions along this curve. Within this plot at low fractal dimension (to the left) we find the fairly demanding cognitive tasks of inventing nonsense words beginning with the letter "r" or ways of entering a house. At high fractal dimension (to the right) we find the more rote tasks such as naming 20 numbers or recalling various foods. Similarly, within the temporal characterization of the Creativity Machine output we find a similar dichotomy between data points falling on the rightmost and leftmost extremes of the plot, where ns/N was maintained constant and equal to 0.06. Leftmost points occurring at a low fractal dimension represent extremes in either s or n (i.e., n = 2 and s = 9 or n = 9 and s = 2). Alternately, the rightmost points on the plot at the high fractal dimension correspond to intermediate values of both s and n (i.e., n = 4 and s = 4.5).

To illustrate this relationship between fractal dimension of the network's output stream and its internal perturbations, I have trained a small feedthrough net on the results of 100 computer experiments to map both n and s to the resulting fractal dimension, D_0. Propagating the n-s array through this trained net I have obtained the plot of Figure 9. There we may observe an asymmetric response between n and s, perhaps indicating the reduced likelihood of smaller perturbations all randomly clustering on the same connection weight to produce an equivalent effect to a large perturbation centered there. Also visible is a central plateau at higher fractal dimension, corresponding to nearly equal, intermediate values of n and s.

Therefore, literally equating the two processes of human cognition and cavitation within the Creativity Machine paradigm, simple tasks such as counting and rote memory recall seem to involve the distribution of many intermediate strength perturbations among many different connection weights of the system. By contrast, more inventive cognitive forays appear to involve the spontaneous appearance of large perturbations localized to just a few connection weights. Admittedly, it will be harder for the cortical or synthetic network to perform internal vector completion on a large, local spike in connection weight perturbation than on smaller distributed disturbances. For a dramatic internal perturbation, the network cannot easily fall into an attractor basin representing a known training exemplar (i.e., a memory). Such a large perturbation can only disrupt and transform the local attractor basin structure to create new and unique attractors. It is these newly formed attractor basins derived from the established memories that now constitute corrupted memories that may or may not be of utilitarian value to connected associative networks.
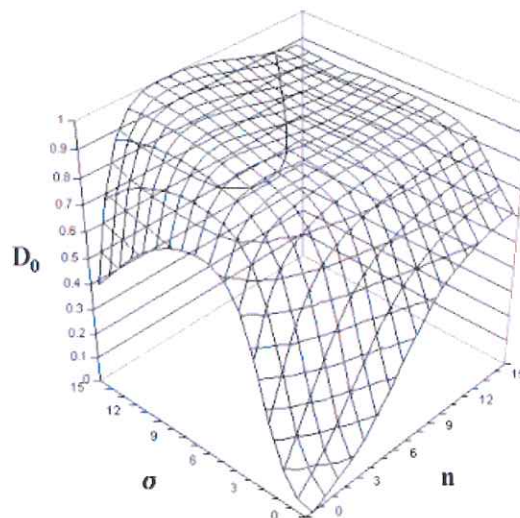

Figure 9. Neural Network Fit to D_0 As a Function of n and s for the Cavitating Artificial Neural Network.

We are accordingly drawn to the conclusion that what the common parlance calls memories and ideas are derivatives of one another that fall at opposite extremes of network perturbations. In short, inventive thought may be no more than internally generated confabulations nucleated upon large synaptic noise spikes, and deemed valuable or interesting by the surrounding cortical networks.

The Apparent Origin of Cognition and Creativity Machine Output in Fractal Brownian Motion

In view of the striking result that both cognitive and Creativity Machine conceptual streams obey the same empirically determined law, it may be worthwhile to develop an ab initio theory, at least for the computational model, to account for the simple, yet all-encompassing relationship contained in Equation 5. Viewed from the perspective of an average neuron embedded within the cavitating IE, the mean activation is essentially performing a random walk, occasionally and sporadically passing through the "all or nothing" activation threshold that toggles the computational neuron into its on or off state. This random walk may be imagined to take place in a series of infinitesimally small steps, each of which is independent of the previous one, thus qualifying the system for analysis via the theory of fBm (Voss, 1988).

A general result from the theory of fractal Brownian motion states that if some time-dependent function $F_H(t)$ is the sum of independent increments or jumps, then the typical change in F, $DF = F(t_2)-F(t_1)$, in the time interval $Dt = t_2 - t_1$, is given by the simple scaling law

$$DF = kDt^H, \qquad (6)$$

where H is in the range $0<H<1$ and k is a unitary dimensional constant. The fractal dimension, D, of the resulting functional trace, $F_H(t)$, is given by the simple relationship,

$$D = 2 - H. \qquad (7)$$

Furthermore, the intersection of this fractal curve with the time axis generates a set of points, known as the "zeroset," with a fractal dimension $D_0=D-1$. We may then recast Equation 1 in terms of the functional trace's zeroset fractal dimension,

$$DF = kDt^{(1-D_0)}, \qquad (8)$$

where the unitary constant k has the dimensions of $sec^{(D_0-1)}$ if Dt is expressed in seconds.

Similarly, if the chaotic net input to the representative mean neuron also performs a random walk, in a series of small independent steps, then the RMS variation in that net input, Dnet, varies as

$$Dnet = kDt^{(1-D_0)}, \qquad (9)$$

where Dt is the observation time and D_0 is the fractal dimension of net input's zeroset. Assuming a bias level of zero to the mean neuron, the zeroset dimension, D_0, takes on the significance of the fractal dimension of mean neuronal on-off transitions. Rearranging Equation 7, we obtain

$$\ln(Dnet /Dt) = - D_0 \ln(kDt). \qquad (10)$$

Assuming that the average activation for all computational neurons is effectively ½ (i.e., the computational neurons may only activate within the range from 0 to 1), and noting that over any half cycle portrayed in Figure 2, the average change in net input is ns/N within a time frame dt/2, the average net input transition rate for a mean neuron is given by ½*2ns/Ndt = ns/Ndt. Substituting this value for Dnet/Dt in Equation 10, we obtain the empirically obtained functional form of Equation 4.

$$\ln(ns/Ndt) = -D_0 \ln(kDt). \qquad (11)$$

Hence for a fixed level of internal network perturbation (ns/Ndt = constant) the product of the mean neuron's output zeroset dimension and the logarithm of the observation time should likewise be constant. We note that the right side of Equation 11 is related to the logarithm of the total number of network transitions, N_0 observed, by the definition of fractal dimension contained in Figure 5. Therefore, Dt may then be thought of as the time required to observe N_0 mean neuron transitions, or an equivalent number of distinct activation transitions within the network as a whole. We note that because this analysis has been fractal, and hence time-scale invariant, any piece of the event stream should yield the same fractal dimension. Hence, the analysis would lead to identical results should there appear some systematic excursion in the midst of the event stream (i.e., the intentional application of inputs to the net) where we might observe a vertical discontinuity in the example fractal traces of Figure 10.

Exponentiation of both sides of Equation 11 and omitting the dimensional constant k leads to

$$Dt^{-D_0} = ns/Ndt, \qquad (12)$$

reproducing Equation 5, with substitution of the optimal cavitation rate $n_s/N dt = 0.19\ \text{sec}^{-1} = r_0$.
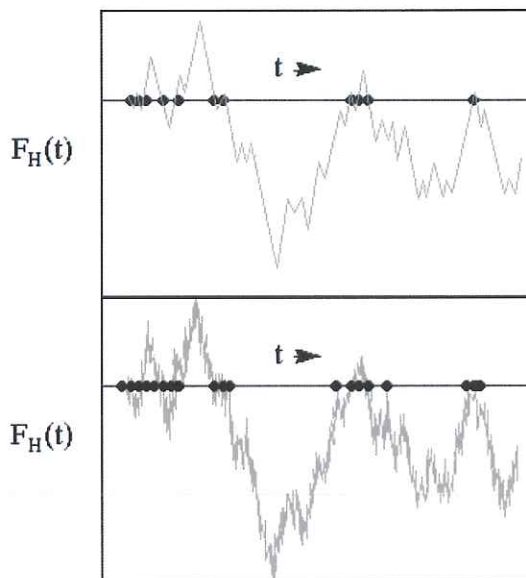

Figure 10. Two Traces of Fractal Brownian Motion. In the upper trace, incremental jumps are scaling as Equation 4, with H = 0.9. Its fractal dimension is therefore 1.1, making it nearly linear. The lower trace corresponds to H = 0.1 and hence a fractal dimension of 1.9, making it nearly a two dimensional object in the Euclidean sense. H=0.5 would correspond to the classic random walk in one dimension. The zersoset is represented as the intersection of each trace with the t-axis. Note the higher the fractal dimension of the trace, the higher the fractal dimensionality of the zeroset with accordingly more intercepts with the time axis..

Returning to the equivalent cognitive result, empirically discovered through the cumulative plot of Figure 8, we must conclude that the cognitive event streams sampled within this study display the signature of fractal Brownian motion by their adherence to Equation 12. This result strongly suggests a dominant mechanism behind all forms of cognition, namely the stochastic perturbation of synaptic connections between biological neurons (or in an equivalent circuit sense, perturbations within the neuron itself). Furthermore, the quantitative applicability of Equation 12 to human cognition, as well as the universality of this result for all artificial neural network structures, indicates that similar levels of network perturbation, $r_0 = 0.19\ \text{sec}^{-1}$, separate regimes of straightforward memory recall from those of novel thought generation within the human cortex.

Conclusions

The rhythm of concept generation within both human test subjects and the Creativity Machine have been shown to be identical for all data gathered within this study. These results substantiate earlier investigations ([Thaler, 1996b](#)) probing the temporal behavior of diverse cavitating neural networks spanning a wide range of sizes, complexities, and connectivities. The remainder of the comparison with human cognition, the singularity and significance of Creativity Machine discoveries and inventions, will always be open to debate, as is the case with any human innovator who must battle against a variety of societal forces (i.e., consensus opinion) and competitive pressures no matter how inherently valuable his or her concepts may be.

Nevertheless, the temporal and fractal equivalence between Creativity Machine concept generation and human cognition is striking, strongly suggesting that stream of consciousness, both mundane and novel, follows the same empirically discoverable laws. The fact that the prosody of cortical concept formation shows the signature of fractal Brownian motion, strongly suggests that stochastic, and perhaps chaotic phenomena within biological neural networks are at the heart of all cognition.

Since the all-neural Creativity Machine demonstrates identical time evolution with human cognition, and because it is capable of producing both incremental and paradigm shift thinking, we may consider this canonical system to be a potential functional model of human cognition. By analogy with the computationally transparent Creativity Machine, rote memory recall appears to be the result of a relatively uniform distribution of small perturbations spread across many connection weights. Alternatively, novel concept formation is tied to the sporadic appearance of relatively large and localized perturbations. Because of their significant effect on the attractor landscape of the network, these larger perturbations may readily alter, merge, and separate specific attractor basins, representing distinct memories and concepts, into modified or perhaps hybridized notions.

Further, because the connections weights within an artificial neural network constitute the statistical rules and correlations that bind together any conceptual space, we may view the process of cavitation as a stochastic experiment within the net in which each of the underlying rules and conventions are randomly softened, singly or in parallel, to produce derivative concepts beyond those experienced within network training. Viewed in the context of hopping perturbations, weight disruptions may sporadically congregate on specific connection weight traces constituting what we would normally consider symbolically representable rules. When such disruptions occur, the singled-out conventions are modified, for better or for worse, as judged by the critic network's response to the emerging concepts. Improvement in the search efficiency of such a system comes in the ability of the policing network to identify and selectively soften those connection traces cumulatively learned by that critic to be essential to the emergence of useful concepts.

In spite of its simplicity, Equation 12 may broadly describe the gamut of human cognition with the perturbation rate of $ns/Ndt = r_0 = 0.19 \, sec^{-1}$, representing a fairly universal constant within neurobiology. Since observed "bubble" formation in cortex takes place on a time scale of roughly 300 msec (Taylor, 1996), such concept nucleating events may be tantamount to the cavitation cycle described in Figure 2. Following this analogy to its conclusion, a similar mean perturbation rate per synapse may apply within neurobiology, perhaps qualifying r_0 as the dividing line between mundane and creative thought. Within the latter creative regime, the random release or diffusion of various neuromodulators and neurohormones could easily provide the intense local perturbations necessary for novel concept formation. Perhaps, this observation accounts in broad sense for the observed correlation between artistic creativity and various neurochemical imbalances such as manic-depressive illness (Jamison). Viewed in this sense, creativity may represent a talent or, alternatively, an unwelcome propensity for the cortex to biochemically 'spike' itself beyond this threshold perturbation level, on demand or otherwise.

Placed on the same continuum of perturbation, all cognition may be viewed as acts of creativity. Even at the lowest levels of synaptic disruption one idea is supplanted by another in a display of low-level originality, as in everyday stream of consciousness, conversation, or movement planning. The noblest invention, scientific discovery, or artistic inspiration lies at the opposite extreme of this spectrum, where the Creativity Machine model implicates large localized perturbations as the nucleating events. Within either of these regimes, consciousness itself may be no more than the spontaneous invention of significance by associative cortical networks to the endless noise-driven activations of their brethren. Our search for an objective truth regarding the basis of creativity and consciousness alike may be blinded by the capacity of such networks to overwhelm and distract us with multiple drafts (i.e., Dennett, 1991) of the actual underlying processes. This dynamical systems approach is an attempt to circumvent an inevitable philosophical cul-de-sac.

References

Dennett, D.C. (1991). Consciousness Explained, Little Brown, Boston.

Jamison, K.R. (1994). Manic Depressive Illness and Creativity, *Scientific American*, 272(2), 62–67.

Kosslyn, S. and Koenig, R. (1992). Wet Mind: The New Cognitive Neuroscience, The Free Press, New York, pp. 241–242.

Mandelbrot, B.B. and van Ness, J.W. (1982). *SIAM Review*, 10(4), 422–437.

Rowe, J. and Partridge, G. (1993). Creativity: A Survey of AI Approaches, *Artificial Intelligence Review*, 7, 43–70.

Rumelhart, D.E., Smolensky, P., McClelland, J.L., and Hinton, G.E. (1986). Schemata and Sequential Thought Processes in PDP Models, In Parallel Distributed Processing, Exploration in the Microstructure of Cognition, Volume 2: Psychological and Biological Models, MIT Press, Cambridge, MA, pp. 7–57.

Taylor, J. (1996). Consciousness: The Relational Mind and It's Emergence, *In World Congress on Neural Networks*, (WCNN'96), Lawrence Erlbaum, Mawah, NJ.

Thaler, S.L. (1995). "Virtual Input Phenomena" Within the Death of a Simple Pattern Associator, *Neural Networks*, 8(1), 55–65.

Thaler, S.L. (1996a). Neural Nets That Create and Discover, *PC AI*, May/June, 16–21.

Thaler, S.L. (1996b). Is Neuronal Chaos the Source of Stream of Consciousness? In Proceedings of the World Congress on Neural Networks, (WCNN'96), Lawrence Erlbaum, Mawah, NJ.

Thaler, S.L. (1996c). A Proposed Symbolism for Network-Implemented Discovery Processes, In Proceedings of the World Congress on Neural Networks, (WCNN'96), Lawrence Erlbaum, Mawah, NJ.

Thaler, S.L. (1996d). Autonomous Materials Discovery Via Spreadsheet-Implemented Neural Network Cascades, Journal of the Minerals, Metals, and Materials Society, *JOM-e*, 49(4) [http://www.tms.org/pubs/journals/JOM/9704/Thaler]

Voss, R. (1988). Fractals in nature: From characterization to simulation, In The Science of Fractal Images, H. Peitgen and D. Saupe, Eds., Springer–Verlag, New York, pp. 21–70.

* U.S. Patent 5,659,666, Device for the Autonomous Generation of Useful Information, Issued 8/19/97